SEQUENCE LISTING

<110> BIER, ETHAN
YU, KWEON

<120> PEPTIDE INHIBITOR OF TGF-BETA GROWTH FACTORS

<130> 041673/2007

<140> US 09/215,569
<141> 1998-12-16

<150> US 60/069,701
<151> 1997-12-16

<160> 12

<170> PatentIn Ver. 2.1

<210> 1
<211> 879
<212> DNA
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 1
atggccaaca agctgaggaa atcgaacgcc atcgaatggg ccagcgccac cggcacagta
60
ccgctcctgg aaaggagctg ctgccacacg gaggacgccg cactggagcc ccaagcgagc
120
aaaaccagcc atagagaaca agccccccatc ctgcgccacc tgagccaact gagccacctg
180
ctcatcatcg ccccactgct gatcgtctgc ttgcgccccgg tgacggaggg ccgccggcat
240
gcgccgctca tgttcgagga gtccgacacg ggcaggcggt ccaaccgacc agcggtcacc
300
gaatgccagt ttggcaaagt tttgcgcgaa ttggggtcca cctggtatgc ggatttgggt
360
ccacccttcg gagttatgta ctgcatcaag tgtgaatgtg tgggcgatacc caagaagcgg
420
cgcatcgttg cacgcgtcca gtgtgcgaat atcaaaaacg agtgcccgcc ggccaaatgc
480
gatgatccca tctcgttgcc cggaaaatgc tgcaagacct gtcccggcga tcgaaacgat
540
acggatgtag ccttggatgt gcccgtgccc aatgaagagg aagagcgcaa catgaaacat
600
tacgctgcgt tgctaacggg ccgcacctcc tatttcctca agggtgagga aatgaagtcc
660
atgtacacca ctacaatcc gcagaatgtg gtggccaccg cccgtttcct gttccacaag
720

aagaatctat actactcctt ctacacctca tcgcgaatcg gtcgtccgcg tgccattcaa
780
ttcgttgatg atgcgggtgt aatcctggag gagcatcaac tggagaccac cttgcgcggc
840
actctcagtg tctatcagaa tgccacgggc aagatctga
879

<210> 2
<211> 292
<212> PRT
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 2

```
Met Ala Asn Lys Leu Arg Lys Ser Asn Ala Ile Glu Trp Ala Thr Ala
  1               5                  10                  15

Thr Gly Thr Val Pro Leu Leu Glu Arg Ser Cys Cys His Ser Glu Asp
             20                  25                  30

Ala Ala Leu Glu Pro Gln Ala Ser Lys Thr Ser His Arg Glu Gln Ala
         35                  40                  45

Pro Ile Leu Arg His Leu Ser Gln Leu Ser His Leu Leu Ile Ile Ala
     50                  55                  60

Gly Leu Leu Ile Val Cys Leu Ala Gly Val Thr Glu Gly Arg Arg His
 65                  70                  75                  80

Ala Pro Leu Met Phe Glu Glu Ser Asp Thr Gly Arg Arg Ser Asn Arg
                 85                  90                  95

Pro Ala Val Thr Glu Cys Gln Phe Gly Lys Val Leu Arg Glu Leu Gly
            100                 105                 110

Ser Thr Trp Tyr Ala Asp Leu Gly Pro Pro Phe Gly Val Met Tyr Cys
        115                 120                 125

Ile Lys Cys Glu Cys Val Ala Ile Pro Lys Lys Arg Arg Ile Val Ala
    130                 135                 140

Arg Val Gln Cys Arg Asn Ile Lys Asn Glu Cys Pro Pro Ala Lys Cys
145                 150                 155                 160

Asp Asp Pro Ile Ser Leu Pro Gly Lys Cys Cys Lys Thr Cys Pro Gly
                165                 170                 175

Asp Arg Asn Asp Thr Asp Val Ala Leu Asp Val Pro Val Pro Asn Glu
            180                 185                 190

Glu Glu Glu Arg Asn Met Lys His Tyr Ala Ala Leu Leu Thr Gly Arg
        195                 200                 205

Thr Ser Tyr Phe Leu Lys Gly Glu Glu Met Lys Ser Met Tyr Thr Thr
    210                 215                 220

Tyr Asn Pro Gln Asn Val Val Ala Thr Ala Arg Phe Leu Phe His Lys
```

225 230 235 240

Lys Asn Leu Tyr Tyr Ser Phe Tyr Thr Ser Ser Arg Ile Gly Arg Pro
245 250 255

Arg Ala Ile Gln Phe Val Asp Asp Ala Gly Val Ile Leu Glu Glu His
260 265 270

Gln Leu Glu Thr Thr Leu Ala Gly Thr Leu Ser Val Tyr Gln Asn Ala
275 280 285

Thr Gly Lys Ile
290

<210> 3
<211> 978
<212> DNA
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 3
atggccaaca agctgaggaa atcgaacgcc atcgaatggg ccacggccac cggcacagta
60
ccgctcctgc aaaggagctg ctgccacagc gaggacgccg cactggagcc ccaagcgagc
120
aaaaccagcc atagagaaca agcccccatc ctgcgccacc tgagccaact gagccacctg
180
ctcatcatcg ccggactgct gatcgtctgc ttggcgggcg tgacggaggg ccgccggcat
240
cgcccgctca tgttcgagga gtccgacacg ggcaggcggt ccaaccgacc agcggtcacc
300
gaatgccagt ttggcaaagt tttgcgcgaa ttggggtcca cctggtatgc ggatttgggt
360
ccacccttcg gagttatgta ctgcatcaag tgtgaatgtg tggcgatacc caagaagcgg
420
cgcatcgttg cacgcgtcca gtgtgcgaat atcaaaaacg agtgcccgcc ggccaaatgc
480
gatgatccca tctcgttgcc cggaaaatgc tgcaagacct gtcccggcga tcgaaacgat
540
acggatgtag ccttggatgt gcccgtgccc aatgaagagg aagagcgcaa catgaaacat
600
tacgctgcgt tgctaacggg ccgcacctcc tatttcctca agggtgagga aatgaagtcc
660
atgtacacca cctacaatcc gcagaatgtg gtggccaccg cccgtttcct gttccacaag
720
aagaatctat actactcctt ctacacctca tcgcgaatcg gtcgtccgcg tgccattcaa
780
ttcgttgatg atgcgggtgt aatcctggag gagcatcaac tggagaccac cttggcgggc
840
actctcagtg tctatcagaa tgccacgggc aagatcggcc gcggctcgag ggtacctcta
900
gaggatcttt gtgaaggaac cttacttctg tggtgtgaca taattggaca aactacctac
960
agagatttaa agctctaa
978

<210> 4
<211> 325
<212> PRT
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 4

```
Met Ala Asn Lys Leu Arg Lys Ser Asn Ala Ile Glu Trp Ala Thr Ala
  1               5                  10                  15

Thr Gly Thr Val Pro Leu Leu Glu Arg Ser Cys Cys His Ser Glu Asp
             20                  25                  30

Ala Ala Leu Glu Pro Gln Ala Ser Lys Thr Ser His Arg Glu Gln Ala
         35                  40                  45

Pro Ile Leu Arg His Leu Ser Gln Leu Ser His Leu Leu Ile Ile Ala
     50                  55                  60

Gly Leu Leu Ile Val Cys Leu Ala Gly Val Thr Glu Gly Arg Arg His
 65                  70                  75                  80

Ala Pro Leu Met Phe Glu Glu Ser Asp Thr Gly Arg Arg Ser Asn Arg
                 85                  90                  95

Pro Ala Val Thr Glu Cys Gln Phe Gly Lys Val Leu Arg Glu Leu Gly
            100                 105                 110

Ser Thr Trp Tyr Ala Asp Leu Gly Pro Pro Phe Gly Val Met Tyr Cys
        115                 120                 125

Ile Lys Cys Glu Cys Val Ala Ile Pro Lys Lys Arg Arg Ile Val Ala
    130                 135                 140

Arg Val Gln Cys Arg Asn Ile Lys Asn Glu Cys Pro Pro Ala Lys Cys
145                 150                 155                 160

Asp Asp Pro Ile Ser Leu Pro Gly Lys Cys Cys Lys Thr Cys Pro Gly
                165                 170                 175

Asp Arg Asn Asp Thr Asp Val Ala Leu Asp Val Pro Val Pro Asn Glu
            180                 185                 190

Glu Glu Glu Arg Asn Met Lys His Tyr Ala Ala Leu Leu Thr Gly Arg
        195                 200                 205

Thr Ser Tyr Phe Leu Lys Gly Glu Glu Met Lys Ser Met Tyr Thr Thr
    210                 215                 220

Tyr Asn Pro Gln Asn Val Val Ala Thr Ala Arg Phe Leu Phe His Lys
225                 230                 235                 240
```

Lys Asn Leu Tyr Tyr Ser Phe Tyr Thr Ser Ser Arg Ile Gly Arg Pro
245 250 255

Arg Ala Ile Gln Phe Val Asp Asp Ala Gly Val Ile Leu Glu Glu His
260 265 270

Gln Leu Glu Thr Thr Leu Ala Gly Thr Leu Ser Val Tyr Gln Asn Ala
275 280 285

Thr Gly Lys Ile Gly Arg Gly Ser Arg Val Pro Leu Glu Asp Leu Cys
290 295 300

Glu Gly Thr Leu Leu Leu Trp Cys Asp Ile Ile Gly Asn Thr Thr Tyr
305 310 315 320

Arg Asp Leu Lys Leu
325

<210> 5
<211> 978
<212> DNA
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 5
atggccaaca agctgaggaa atcgaacgcc atcgaatggg ccacggccac cggcacagta
60
ccgctcctgg aaaggagctg ctgccacagc gaggacgccg cactggagcc ccaagcgagc
120
aaaaccagcc atagagaaca agccccccatc ctgcgccacc tgagccaact gagccacctg
180
ctcatcatcg ccggactgct gatcgtctgc ttggcgggcg tgacggaggg ccgccggcat
240
gcgccgctca tgttcgagga gtccgacacg ggcaggcggt ccaaccgacc agcggtcacc
300
gaatgccagt ttggcaaagt tttgcgcgaa ttggggtcca cgcagtatgc ggatttgggt
360
ccaccctteg gagttatgta ctgcatcaag tgtgaatgtg tggcgatacc caagaagcgg
420
cgcatcgttg cacgcgtcca gtgtcgcaat atcaaaaacg agtgcccgcc ggccaaatgc
480
gatgatccca tctcgttgcc cggaaaatgc tgcaagacct gtcccggcga tcgaaacgat
540
acggatgtag ccttggatgt gcccgtgccc aatgaagagg aagagcgcaa catgaaacat
600
tacgctgcgt tgctaacggg ccgcacctcc tatttcctca agggtgagga aatgaagtcc
660
atgtacacca cctacaatcc gcagaatgtg gtggccaccg cccgtttcct gttccacaag
720
aagaatctat actactcctt ctacacctca tcgcgaatcg gtcgtccgcg tgccattcaa
780
ttcgttgatg atgcgggtgt aatcctggag gagcatcaac tggagaccac cttggcgggc
840
actctcagtg tctatcagaa tgccacgggc aagatcggcc gcggctcgag ggtacctcta
900
gaggatcttt gtgaaggaac cttacttctg tggtgtgaca taattggaca aactacctac
960
agagatttaa agctctaa
978

<210> 6
<211> 325
<212> PRT
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 6

```
Met Ala Asn Lys Leu Arg Lys Ser Asn Ala Ile Glu Trp Ala Thr Ala
  1               5                  10                  15

Thr Gly Thr Val Pro Leu Leu Glu Arg Ser Cys Cys His Ser Glu Asp
             20                  25                  30

Ala Ala Leu Glu Pro Gln Ala Ser Lys Thr Ser His Arg Glu Gln Ala
         35                  40                  45

Pro Ile Leu Arg His Leu Ser Gln Leu Ser His Leu Leu Ile Ile Ala
     50                  55                  60

Gly Leu Leu Ile Val Cys Leu Ala Gly Val Thr Glu Gly Arg Arg His
 65                  70                  75                  80

Ala Pro Leu Met Phe Glu Glu Ser Asp Thr Gly Arg Arg Ser Asn Arg
                 85                  90                  95

Pro Ala Val Thr Glu Cys Gln Phe Gly Lys Val Leu Arg Glu Leu Gly
            100                 105                 110

Ser Thr Ala Tyr Ala Asp Leu Gly Pro Pro Phe Gly Val Met Tyr Cys
        115                 120                 125

Ile Lys Cys Glu Cys Val Ala Ile Pro Lys Lys Arg Arg Ile Val Ala
    130                 135                 140

Arg Val Gln Cys Arg Asn Ile Lys Asn Glu Cys Pro Pro Ala Lys Cys
145                 150                 155                 160

Asp Asp Pro Ile Ser Leu Pro Gly Lys Cys Cys Lys Thr Cys Pro Gly
                165                 170                 175

Asp Arg Asn Asp Thr Asp Val Ala Leu Asp Val Pro Asp Pro Asn Glu
            180                 185                 190

Glu Glu Glu Arg Asn Met Lys His Tyr Ala Ala Leu Leu Thr Gly Arg
        195                 200                 205

Thr Ser Tyr Phe Leu Lys Gly Glu Glu Met Lys Ser Met Tyr Thr Thr
    210                 215                 220

Tyr Asn Pro Gln Asn Val Val Ala Thr Ala Arg Phe Leu Phe His Lys
```

225 230 235 240

Lys Asn Leu Tyr Tyr Ser Phe Tyr Thr Ser Ser Arg Ile Gly Arg Pro
245 250 255

Arg Ala Ile Gln Phe Val Asp Asp Ala Gly Val Ile Leu Glu Glu His
260 265 270

Lys Leu Glu Thr Thr Leu Ala Gly Thr Leu Ser Val Tyr Gln Asn Ala
275 280 285

Thr Gly Lys Ile Gly Arg Gly Ser Arg Val Pro Leu Glu Asp Leu Cys
290 295 300

Glu Gly Thr Leu Leu Leu Trp Cys Asp Ile Ile Gly Asn Thr Thr Tyr
305 310 315 320

Arg Asp Leu Lys Leu
325

<210> 7
<211> 1041
<212> DNA
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 7
atggccaaca agctgaggaa atcgaacgcc atcgaatggg ccacggccac cggcacagta
60
ccgctcctgg aaaggagctg ctgccacagc gaggacgccg cactggagcc ccaagcgagc
120
aaaaccagcc atagagaaca agcccccatc ctgcgccacc tgagccaact gagccacctg
180
ctcatcatcg ccggactgct gatcgtctgc ttggcgggcg tgacggaggg ccgccggcat
240
gcgccgctca tgttcgagga gtccgacacg ggcaggcggt ccaaccgacc agcggtcacc
300
gaatgccagt ttggcaaagt tttgcgcgaa ttggggtcca cctggtatgc ggatttgggt
360
ccacccttcg gagttatgta ctgcatcaag tgtgaatgtg tggcgatacc caagaagcgg
420
cgcatcgttg cacgcgtcca gtgtcgcaat atcaaaaacg agtgcccgcc ggccaaatgc
480
gatgatccca tctcgttgcc cggaaaatgc tgcaagacct gtcccggcga tcgaaacgat
540
acggatgtag ccttggatgt gcccgtgccc aatgaagagg aagagcgcaa catgaaacat
600
tacgctgcgt tgctaacggg cgcgacctcc tatttcctca agggtgagga aatgaagtcc
660
atgtacacca cctacaatcc gcagaatgtg gtggccaccg cccgtttcct gttccacaag
720
aagaatctat actactcctt ctacacctca tcgcgaatcg gtcgtccgcg tgccattcaa
780
ttcgttgatg atgcgggtgt aatcctggag gagcatcaac tggagaccac cttggcgggc
840
actctcagtg tctatcagaa tgccacgggc aagatcggcc gcggctcgag gcagcgcggc
900
cgcatctttt acccatacga tgttcctgac tatgcgggct atcctatga cgtcccggac
960
tatgcaggat cctatccata tgacgttcca gattacgctg ctcagtgcgg ccgcgattat
1020
agggacgacg acgacaaatc a
1041

<210> 8
<211> 346
<212> PRT
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 8

```
Met Ala Asn Lys Leu Arg Lys Ser Asn Ala Ile Glu Trp Ala Thr Ala
  1               5                  10                  15

Thr Gly Thr Val Pro Leu Leu Glu Arg Ser Cys Cys His Ser Glu Asp
             20                  25                  30

Ala Ala Leu Glu Pro Gln Ala Ser Lys Thr Ser His Arg Glu Gln Ala
         35                  40                  45

Pro Ile Leu Arg His Leu Ser Gln Leu Ser His Leu Leu Ile Ile Ala
     50                  55                  60

Gly Leu Leu Ile Val Cys Leu Ala Gly Val Thr Glu Gly Arg Arg His
 65                  70                  75                  80

Ala Pro Leu Met Phe Glu Glu Ser Asp Thr Gly Arg Arg Ser Met Arg
                 85                  90                  95

Pro Ala Val Thr Glu Cys Gln Phe Gly Lys Val Leu Arg Glu Leu Gly
            100                 105                 110

Ser Thr Trp Tyr Ala Asp Leu Gly Pro Pro Phe Gly Val Met Tyr Cys
        115                 120                 125

Ile Lys Cys Glu Cys Val Ala Ile Pro Lys Lys Arg Arg Ile Val Ala
    130                 135                 140

Arg Val Gln Cys Arg Asn Ile Lys Asn Glu Cys Pro Pro Ala Lys Cys
145                 150                 155                 160

Asp Asp Pro Ile Ser Leu Pro Gly Lys Cys Cys Lys Thr Cys Pro Gly
                165                 170                 175

Asp Arg Asn Asp Thr Asp Val Ala Leu Asp Val Pro Val Pro Asn Glu
            180                 185                 190

Glu Glu Glu Arg Asn Met Lys His Tyr Ala Ala Leu Leu Thr Gly Arg
        195                 200                 205

Thr Ser Tyr Phe Leu Lys Gly Glu Glu Met Lys Ser Met Tyr Thr Thr
    210                 215                 220

Tyr Asn Pro Gln Asn Val Val Ala Thr Ala Arg Phe Leu Phe His Lys
```

225 230 235 240

Lys Asn Leu Tyr Tyr Ser Phe Tyr Thr Ser Ser Arg Ile Gly Arg Pro
245 250 255

Arg Ala Ile Gln Phe Val Asp Asp Ala Gly Val Ile Leu Glu Glu His
260 265 270

Gln Leu Glu Thr Thr Leu Ala Gly Thr Leu Ser Val Tyr Gln Met Ala
275 280 285

Thr Gly Lys Ile Gly Arg Gly Ser Arg Asn Arg Gly Arg Ile Phe Tyr
290 295 300

Pro Tyr Asp Val Pro Asp Tyr Ala Gly Tyr Pro Tyr Asp Val Pro Asp
305 310 315 320

Tyr Ala Gly Ser Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ala Asn Cys
325 330 335

Gly Arg Asp Tyr Lys Asp Asp Asp Asp Lys
340 345

<210> 9
<211> 213
<212> PRT
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 9
Met Asp His Ser Gln Cys Leu Val Thr Ile Tyr Ala Leu Met Val Phe
1 5 10 15

Leu Gly Leu Arg Ile Asp Gln Gly Gly Cys Gln His Tyr Leu His Ile
20 25 30

Arg Pro Ala Pro Ser Glu Asn Leu Pro Leu Val Asp Leu Ile Glu His
35 40 45

Pro Asp Pro Ile Tyr Asp Pro Lys Glu Lys Asp Glu Leu Asn Glu Thr
50 55 60

Leu Leu Arg Thr Leu Met Val Gly His Phe Asp Pro Ile Leu Pro Glu
65 70 75 80

Glu Arg Leu Gly Val Glu Asp Leu Gly Glu Leu Asp Leu Leu Leu Arg
85 90 95

Gln Lys Pro Ser Gly Ala Met Pro Ala Glu Ile Lys Gly Leu Glu Phe
100 105 110

```
Tyr Glu Gly Leu Gln Ser Lys Lys His Arg Leu Ser Lys Lys Leu Arg
        115                 120                 125

Arg Lys Leu Gln Met Trp Leu Trp Ser Gln Thr Phe Cys Pro Val Leu
    130                 135                 140

Tyr Thr Trp Asn Asp Leu Gly Arg Tyr Val Lys Val Gly Ser Cys Tyr
145                 150                 155                 160

Ser Lys Arg Ser Cys Ser Val Pro Glu Gly Met Val Cys Lys Ala Ala
                165                 170                 175

Lys Ser Met His Leu Thr Ile Leu Arg Trp Arg Cys Gln Arg Arg Val
            180                 185                 190

Gln Gln Lys Cys Ala Trp Ile Thr Ile Gln Tyr Pro Val Ile Ser Glu
        195                 200                 205

Cys Lys Cys Ser Cys
    210
```

<210> 10
<211> 53
<212> PRT
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 10

```
Val Asp Leu Ile Glu His Pro Asp Pro Ile Tyr Asp Pro Lys Glu Lys
  1               5                  10                  15

Asp Leu Asn Glu Thr Leu Leu Arg Thr Leu Met Val Gly His Phe Asp
            20                  25                  30

Pro Ile Leu Pro Glu Glu Arg Leu Gly Val Glu Asp Leu Gly Glu Leu
        35                  40                  45

Asp Leu Leu Leu Arg
    50
```

<210> 11
<211> 4892
<212> DNA
<213> Short Gastrulation Gene (SOG)

<400> 11
atattagatg gacatgcata ataattattc atgtaactat gtgattttca ttttacacga
60
ggtgtcagtc agaatttaaa attcttaaaa attgcaatca cggtctattg tacatattta
120
tgtatgcgat ctcattatta ttattattat tatttgataa tatattagca gctgttcaaa
180
tcatcaacaa gtataaaatt gaaaatgtta aaaattgaag cagtcttggg tttgcctgca
240
acatgttgct gccgatcgtt agatgttgct gctgcatgtt gccgctgcat gttgccgttg
300
catgttgccg ccgtttggca actttataaa cacggagcgg attcagtcgt tcaggttcag
360
tctctcttga attgcgtggg attgcacatc ggtcgttcgg ctttttgggt tcggcattta
420
gagagatacg atacgatccg atacgatccg atccagtaca aaaatcaaat tcaaaccgca
480
ctcccgatcc ggtcgccatc atatacacgg cggctcgcac cgcagctatc tagtagataa
540
aaagtcgccg agaacgcagt acgcatatag cgaaacgcca aaaaaataaa aaaaaaagtc
600
gcgttgtcgt aatcgctcct catacgagat cgactctatt ttccagagca agctaaacac
660
actagtgcta aaccataact atatatctaa ctaaggaaaa caaagtctcg aaactgaaaa
720
cgaaaagcgc aaatttatgc agccgctaaa taaaaacaga aaccaaaaca taaaacacac
780
tatacaaatc atacaaaaca aaaacagcga atcaaatagt ataaaaaaaa aataaataaa
840
tgaggaataa aaaaacaggc aaatagaata aatccaataa atcggcgcgc gaaactccgc
900
tgtgttatct aatctgcaag agaagtacaa gaatcgggta tagaatcggc tctatactat
960
atctatacac ctgatatatc tatatccatt gtgtgtgcca gtgtgtgcgt ggcgaccttt
1020
gttttatat attttttgtt gttgttcata ctgtgaaacg tgcttttac aagccggtcg
1080
ttcaaaatac aaaatactac aaatcaaatc aaatcatata cacatacata tcagtaaaaa
1140
caaaacaaaa aacacataaa catggccaac aagctgagga aatcgaacgc catcgaatgg
1200
gccacggcca ccggcacagt accgctcctg gaaaggagct gctgccacag cgaggacgcc
1260

gcactggagc cccaagcgag caaaaccagc catagagaac aagcccccat cctgcgccac
1320
ctgagccaac tgagccacct gctcatcatc gccggactgc tgatcgtctg cttggcgggc
1380
gtgacggagg gccgccggca tcgcccgctc atgttcgagg agtccgacac gggcaggcgg
1440
tccaaccgac cagcggtcac cgaatgccag tttggcaaag ttttgcgcga attggggtcc
1500
acctggtatg cggatttggg tccacccttc ggagttatgt actgcatcaa gtgtgaatgt
1560
gtggcgatac ccaagaagcg gcgcatcgtt gcacgcgtcc agtgtcgcaa tatcaaaaac
1620
gagtgcccgc cggccaaatg cgatgatccc atctcgttgc ccggaaaatg ctgcaagacc
1680
tgtcccggcg atcgaaacga tacggatgta gccttggatg tgcccgtgcc caatgaagag
1740
gaagagcgca acatgaaaca ttacgctgcg ttggtaacgg gccgcacctc ctatttcctc
1800
aagggtgagg aaatgaagtc catgtacacc acctacaatc cgcagaatgt ggtggccacc
1860
gcccgtttcc tgttccacaa gaagaatcta tactactcct tctacacctc atcgcgaatc
1920
ggtcgtccgc gtgccattca attcgttgat gatgcgggtg taatcctgga ggagcatcaa
1980
ctggagacca ccttggcggg cactctcagt gtctatcaga atgccacggg caagatctgc
2040
ggtgtctggc gacgagttcc acgtgattac aagcgcatcc tggcggacga tgctctccat
2100
gttgtcctcc tctggggcaa caaacagcag gccgagttgg ctctggccgg aaaggtggcc
2160
aaatacacgg ccctgcagac ggagttgttc agttcgctac tggaggcacc acttcccgat
2220
ggcaaaacgg atccccagct ggccggagcc ggtggcacag cgatcgtgtc caccagcagc
2280
ggtgccgcct catcgatgca tctcaccctg gtcttcaatg gtgtctttgg tgccgaggag
2340
tacgccgatg cagcactgag tgtgaaaatt gagctggcag aacggaagga ggtgatcttc
2400
gatgagattc cacgtgtgcg caaaccctct gccgagatca atgtcctgga gctgtcgtcg
2460
cccatttcca tacagaatct tgactgatg tcgcgtggca aactcctgct gaccgtggag
2520
tccaagaagt acccacatct gcgcatccag ggacacatcg tgaccgagcc cagctgcgaa
2580
atcttccaga ccctgctggc gccgcacagt gccgaatcct cgaccaagag cagcggtttg
2640
gcgtgggtct acttgaacac cgatggatct ctggcctaca acatcgaaac ggagcacgtg
2700
aacacccggg ataggcccaa catcagtttg attgaggagc agggcaagcg gaaggccaag
2760

ctggaggatc tgacgccgag cttcaacttc aaccaggcca ttggtagtgt ggagaagttg
2820
ggtcccaagg tcctcgagtc gctgtatgcc ggcgaactgg gcgttaatgt ggccaccgag
2880
catgagacga gcctgatccg tggccgccta gtgccccgtc cagtggccga tgctcgggac
2940
tcggcggaac ccattctgct gaagcgacag gagcacacgg atgcacagaa tccacatgcc
3000
gtcggcatgg cctggatgtc cattgacaac gagtgcaatc tgcactacga ggtgacgctc
3060
aacggtgtgc ccgcccagga tctgcagctg tatctggagg agaagcccat cgaggcgatt
3120
ggagcgccag tgacgaggaa attgctcgag gaattcaacg gctcctatct ggaaggcttc
3180
ttcctcagca tgccatccgc cgaactgatc aagctggaga tgagcgtctg ctatctggag
3240
gtccattcca agcactccaa acagcttctg ctgcgcggca aactgaagag caccaaggtg
3300
ccgggtcact gcttcccccgt ctatacggac aacaatgttc ccgtgcctgg cgaccacaat
3360
gataaccatt tggtgaacgg agagaccaag tgctttcact ccggacgctt ctacaacgaa
3420
tcggagcatt ggcgcagtgc ccaggattcc tgtcagatgt gcgcctgttt gcgtggccaa
3480
cacagttgcg aggtcatcaa gtgtccgcct ctcaagtgca agtccacgga gcaactgctt
3540
cagagtgatg gtgaatgctg tcccagctgt gtgcccaaga aggaggccgc cgactattca
3600
gcgcaatcct cgccagccac caatgccacc gatttgctgc aacagcgacg cggctgcgcg
3660
ctgggcgagc agttccatcc cgccggtgcc agttggcatc cattcctgcc gcccaatggc
3720
ttcgatacct gcaccacctg cagctgcgat ccgctgaccc tcgagattcg ctgtcccccgg
3780
ctcgtctgcc cgccgttgca gtgcagcgag aagttggcct atcgtccaga caagaaggca
3840
tgctgcaaga tctgtccgga gggcaagcag agcagttcca atggacacaa gacgacgccg
3900
aacaatccca atgtgctgca ggatcaggcc atgcagcgat cgccgagtca cagtgccgag
3960
gaggttctgg ccaacggcgg atgcaaggtg gtcaacaagg tgtacgagaa cggccaggag
4020
tggcatccga tcctgatgtc ccacggcgag cagaagtgca tcaagtgccg ctgcaaggac
4080
tccaaggtga actgcgatgc caagcgctgc tccgctcca cgtgccagca gcagacacgc
4140
gtgaccagca aacggcgtct gttcgagaaa ccggacgcag ctgctccggc catcgatgac
4200
ttctgctcca ccgactgccg gagatcgagg cgccaccaca agaggcagcc gcatcatcag
4260

cagcgatcct ccagctgagc ggctccacgt gacggatggg atcccaatcc agtatcagat
4320
ccttggcggc aggggagcga accaatcact cactcactca ccaccactca gtgtactcag
4380
tgtgcaccac ccaaacacac acacacacac acacacacaa ccacacaaca ctcacaccca
4440
catctacaca gacacacaga cagccacaaa agcgaacgcg cacacagact tgtgcaagga
4500
gttgcataga tcgttgttgc tatcttatca tgtggcagca atgagaactt gtattatata
4560
tatgaatcac ggaggagaaa acgtaggaga gaaatctcac aaaaaatata tatatcttat
4620
ggaggaaaac ggtagtaata gagagagaga gagagggaag gagagagtct aatgagatcc
4680
ttggaaaagg acattaaaac cagtgcagtt tgctttaaat tctccagcgc agaattttct
4740
attgaaagca ttttctgaat ttcttttcgc agttacccca cccgtgtaac ccaatcccct
4800
cccctcccca accaacaaac acccaaaaaa aaaaactaaa aacattaaaa tacaatttta
4860
atttattaca aaacaaaaac aaaaaaaaaa aa
4892

<210> 12
<211> 1038
<212> PRT
<213> Drosophila Short Gastrulation Gene (SOG)

<400> 12

```
Met Ala Asn Lys Leu Arg Lys Ser Asn Ala Ile Glu Trp Ala Thr Ala
  1               5                  10                  15

Thr Gly Thr Val Pro Leu Leu Glu Arg Ser Cys Cys His Ser Glu Asp
             20                  25                  30

Ala Ala Leu Glu Pro Gln Ala Ser Lys Thr Ser His Arg Glu Gln Ala
         35                  40                  45

Pro Ile Leu Arg His Leu Ser Gln Leu Ser His Leu Leu Ile Ile Ala
     50                  55                  60

Gly Leu Leu Ile Val Cys Leu Ala Gly Val Thr Glu Gly Arg Arg His
 65                  70                  75                  80

Ala Pro Leu Met Phe Glu Glu Ser Asp Thr Gly Arg Arg Ser Asn Arg
                 85                  90                  95

Pro Ala Val Thr Glu Cys Gln Phe Gly Lys Val Leu Arg Glu Leu Gly
            100                 105                 110

Ser Thr Trp Tyr Ala Asp Leu Gly Pro Pro Phe Gly Val Met Tyr Cys
        115                 120                 125

Ile Lys Cys Glu Cys Val Ala Ile Pro Lys Lys Arg Arg Ile Val Ala
    130                 135                 140

Arg Val Gln Cys Arg Asn Ile Lys Asn Glu Cys Pro Pro Ala Lys Cys
145                 150                 155                 160

Asp Asp Pro Ile Ser Leu Pro Gly Lys Cys Cys Lys Thr Cys Pro Gly
                165                 170                 175

Asp Arg Asn Asp Thr Asp Val Ala Leu Asp Val Pro Val Pro Asn Glu
            180                 185                 190

Glu Glu Glu Arg Asn Met Lys His Tyr Ala Ala Leu Leu Thr Gly Arg
        195                 200                 205

Thr Ser Tyr Phe Leu Lys Gly Glu Glu Met Lys Ser Met Tyr Thr Thr
    210                 215                 220

Tyr Asn Pro Gln Asn Val Val Ala Thr Ala Arg Phe Leu Phe His Lys
```

225 230 235 240

Lys Asn Leu Tyr Tyr Ser Phe Tyr Thr Ser Ser Arg Ile Gly Arg Pro
245 250 255

Arg Ala Ile Gln Phe Val Asp Asp Ala Gly Asx Ile Leu Glu Glu His
260 265 270

Gln Leu Glu Thr Thr Leu Ala Gly Thr Leu Ser Val Tyr Gln Asn Ala
275 280 285

Thr Gly Lys Ile Cys Gly Val Trp Arg Arg Val Pro Arg Asp Tyr Lys
290 295 300

Arg Ile Leu Arg Asp Asp Arg Leu His Val Val Leu Leu Trp Gly Asn
305 310 315 320

Lys Gln Gln Ala Glu Leu Ala Leu Ala Gly Lys Val Ala Lys Tyr Thr
325 330 335

Ala Leu Gln Thr Glu Leu Phe Ser Ser Leu Leu Glu Ala Pro Leu Pro
340 345 350

Asp Gly Lys Thr Asp Pro Gln Leu Ala Gly Ala Gly Gly Thr Ala Ile
355 360 365

Val Ser Thr Ser Ser Gly Ala Ala Ser Ser Met His Leu Thr Leu Val
370 375 380

Phe Asn Gly Val Phe Gly Ala Glu Glu Tyr Ala Asp Ala Ala Leu Ser
385 390 395 400

Val Lys Ile Glu Leu Ala Glu Arg Lys Glu Val Ile Phe Asp Glu Ile
405 410 415

Pro Arg Val Arg Lys Pro Ser Ala Glu Ile Asn Val Leu Glu Leu Ser
420 425 430

Ser Pro Ile Ser Ile Gln Asn Leu Arg Leu Met Ser Arg Gly Lys Leu
435 440 445

Leu Leu Thr Val Glu Ser Lys Lys Tyr Pro His Leu Arg Ile Gln Gly
450 455 460

His Ile Val Thr Arg Ala Ser Cys Glu Ile Phe Gln Thr Leu Leu Ala
465 470 475 480

Pro His Ser Ala Glu Ser Ser Thr Lys Ser Ser Gly Leu Ala Trp Val
485 490 495

Tyr Leu Asn Thr Asp Gly Ser Leu Ala Tyr Asn Ile Glu Thr Glu His
500 505 510

Val Asn Thr Arg Asp Arg Pro Asn Ile Ser Leu Ile Glu Glu Gln Gly
515 520 525

Lys Arg Lys Ala Lys Leu Glu Asp Leu Thr Pro Ser Phe Asn Phe Asn
530 535 540

Gln Ala Ile Gly Ser Val Glu Lys Leu Gly Pro Lys Val Leu Glu Ser
545 550 555 560

Leu Tyr Ala Gly Glu Leu Gly Val Asn Val Ala Thr Glu His Glu Thr
565 570 575

Ser Leu Ile Arg Gly Arg Leu Val Pro Arg Pro Val Ala Asp Ala Arg
580 585 590

Asp Ser Ala Glu Pro Ile Leu Leu Lys Arg Gln Glu His Thr Asp Ala
595 600 605

Gln Asn Pro His Ala Val Gly Met Ala Trp Met Ser Ile Asp Asn Glu
610 615 620

Cys Asn Leu His Tyr Glu Val Thr Leu Asn Gly Val Pro Ala Gln Asp
625 630 635 640

Leu Gln Leu Tyr Leu Glu Glu Lys Pro Ile Glu Ala Ile Gly Ala Pro
645 650 655

Val Thr Arg Lys Leu Leu Glu Glu Phe Asn Gly Ser Tyr Leu Glu Gly
660 665 670

Phe Phe Leu Ser Met Pro Ser Ala Glu Leu Ile Lys Leu Glu Met Ser
675 680 685

Val Cys Tyr Leu Glu Val His Ser Lys Met Ser Lys Cys Leu Leu Leu
690 695 700

Arg Gly Lys Leu Lys Ser Thr Lys Val Pro Gly His Cys Phe Pro Val
705 710 715 720

Tyr Thr Asp Asn Asn Val Pro Val Pro Gly Asp His Asn Asp Asn His
725 730 735

Leu Val Asn Gly Glu Thr Lys Cys Phe His Ser Gly Arg Phe Tyr Asn
740 745 750

Glu Ser Glu Gln Trp Arg Ser Ala Gln Asp Ser Cys Gln Met Cys Ala
755 760 765

Cys Leu Arg Gly Gln Ser Ser Cys Glu Val Ile Lys Cys Pro Ala Leu
770 775 780

Lys Cys Lys Ser Thr Glu Gln Leu Leu Gln Arg Asp Gly Glu Cys Cys
785 790 795 800

Pro Ser Cys Val Pro Lys Lys Glu Ala Ala Asp Tyr Ser Ala Gln Ser
805 810 815

Ser Pro Ala Thr Asn Ala Thr Asp Leu Leu Gln Gln Arg Arg Gly Cys
820 825 830

Arg Leu Gly Glu Gln Phe His Pro Ala Gly Ala Ser Trp His Pro Phe
835 840 845

Leu Pro Pro Asn Gly Phe Asp Thr Cys Thr Thr Cys Ser Cys Asp Pro
850 855 860

Leu Thr Leu Glu Ile Arg Cys Pro Arg Leu Val Cys Pro Pro Leu Gln
865 870 875 880

Cys Ser Glu Lys Leu Ala Tyr Pro Pro Asp Lys Lys Ala Cys Cys Lys
885 890 895

Ile Cys Pro Glu Gly Tyr Gln Ser Ser Ser Asn Gly His Lys Thr Thr
900 905 910

Pro Asn Asn Pro Asn Val Leu Gln Asp Gln Ala Met Gln Arg Ser Pro
915 920 925

Ser His Ser Ala Glu Glu Val Leu Ala Asn Gly Gly Cys Lys Val Val
930 935 940

Asn Lys Val Tyr Glu Asn Gly Gln Glu Trp His Pro Ile Leu Met Ser
945 950 955 960

His Gly Glu Gln Lys Cys Ile Lys Cys Arg Cys Lys Asp Ser Lys Val
965 970 975

Asn Cys Asp Ala Lys Arg Cys Ser Arg Ser Thr Cys Gln Gln Gln Thr
980 985 990

Arg Val Thr Ser Lys Arg Arg Leu Phe Glu Lys Pro Asp Ala Ala Ala
995 1000 1005

Pro Ala Ile Asp Glu Phe Cys Ser Thr Gln Cys Arg Arg Ser Arg Arg
1010 1015 1020

His His Lys Arg Gln Pro His His Gln Gln Arg Ser Ser Ser
1025 1030 1035